Controlled acute and follow-up trial of cognitive therapy and pharmacotherapy in out-patients with recurrent depression

I.-M. BLACKBURN and R. G. MOORE

Background We report a randomised controlled trial, in both the acute and maintenance stage of treatment, in 75 out-patients with recurrent major depression.

Method Patients were allocated to three groups: 16 weeks of acute treatment and two years' maintenance treatment in the following way: antidepressants and maintenance antidepressants; cognitive therapy and maintenance cognitive therapy; antidepressants and maintenance cognitive therapy. Both completers' and end-point data were analysed.

Results In the acute phase of treatment, all patients improved significantly and there was no significant difference among treatments, or in the pattern of improvement over time. In the maintenance stage of treatment, patients kept improving over time in all three groups and there was no significant difference among treatments. Cognitive therapy was consistently superior to medication.

Conclusions The results indicate that maintenance cognitive therapy has a similar prophylactic effect to maintenance medication and is a viable option for maintenance after acute treatment with medication in recurrent depression.

Since the seminal controlled trial of nearly 20 years ago comparing the relative efficacy of cognitive therapy (CT) and imipramine in out-patients suffering from major unipolar depression (Rush *et al*, 1977), at least 14 comparable studies have been published comparing CT and antidepressant medication singly or in combination in the acute treatment of out-patients. Six follow-up studies of one year to two years have also been published. A full review of all these studies can be found in several texts (e.g. Blackburn & Twaddle, 1996).

The two most recent controlled studies, the National Institute of Mental Health (NIMH) Treatment of Depression Collaborative Research Program (TDCRP) (Elkin *et al*, 1989; Shea *et al*, 1992) and the Minnesota study of Hollon and colleagues (Hollon *et al*, 1992; Evans *et al*, 1992) report on short- and long-term outcome in the same samples and will be used in the discussion as points of reference.

The results reported in those studies are consistent with the results of other studies comparing the short- and long-term effect of cognitive therapy with that of antidepressant medication, in major unipolar depression. These can be summarised as: (a) cognitive therapy is at least as effective as antidepressant medication; (b) the relative efficacy of cognitive therapy and medication for more severe depression is controvertible (the NIMH study (Elkin *et al*, 1995) reported that medication with clinical management was superior to CT for more severe depression (HRSD score $\geqslant 20$ and Global Assessment Score (GAS; Endicott *et al*, 1976) of $\leqslant 50$); this finding was not replicated by Hollon *et al* (1992), McLean & Taylor (1992) and Thase *et al* (1991)); (c) the combination of cognitive therapy and antidepressants may be somewhat superior to either treatment on its own; and (d) the long-term or prophylactic effect of CT has been demonstrated in both the NIMH and Minnesota Studies and in other previous studies, although all studies have been naturalistic or partially controlled for their medication groups.

This study aims to complement previous results in several novel and important ways. First, treatments during both the acute outcome and follow-up stages of treatment were controlled. To date, all follow-up studies of CT have been naturalistic, with the exception of Evans *et al* (1992) and Blackburn *et al* (1986) which were partially controlled for the medication group. Second, patients entering the study were selected for high risk of recurrence of depression in that all had suffered at least one previous episode of depression. Previous studies have included both first and recurrent episodes. Third, the efficacy of CT as maintenance treatment after acute treatment with medication was tested. This was of particular interest as the prophylactic effect of antidepressants prescribed long-term is well documented, while CT delivered as a maintenance treatment has never been tested. Since antidepressants are generally cheaper to administer than weekly CT, if CT is found to be as effective or more effective than antidepressants when delivered less frequently as continuation treatment, this would make CT more economically viable and possibly more acceptable to patients at risk who do not wish to take medication long-term.

The main objective of the study was to examine the role of CT in the prevention of relapse and recurrence of depression in a fully controlled design which took into consideration treatment received during the acute phase of the illness. Greenhouse *et al* (1991) recommended this procedure to avoid selection bias in the maintenance phase. The null hypotheses were:

(a) In the acute treatment of non-psychotic recurrent major unipolar depression in out-patients, CT will be as effective as antidepressant medication.

(b) In long-term outcome (two years) of recurrent major depression, maintenance CT, after acute treatment with CT, will be as effective as continuation treatment with antidepressants after acute treatment with antidepressants.

(c) Switching from antidepressants to CT during follow-up will be feasible and as effective on follow-up as maintenance medication.

METHOD

Recruitment and screening

Patients were recruited from out-patient referrals to consultants in a large teaching

psychiatric hospital, and from two general practices. All potential referrals were screened by two evaluators, trained research psychologists, who administered the Schedule for Affective Disorders and Schizophrenia (SADS; Endicott & Spitzer, 1978), the Hamilton Rating Scale for Depression (HRSD–17 item; Hamilton, 1960; modified version, as used in the NIMH study, Elkin *et al*, 1989), and took a full psychiatric history. The criteria for entering the study were: age 18–65, a diagnosis of primary major unipolar depression (Research Diagnostic Criteria, RDC; Spitzer *et al*, 1978), non-psychotic, and a score of at least 16 on the HRSD. The current episode had to be at least a second episode of major depression. Patients with another primary axis I disorder, organic brain damage, a history of bipolar illness, alcohol or drug misuse, or who could not be prescribed antidepressant medication for medical reasons, were excluded. Patients were classified as endogenous or not endogenous on two sets of criteria, the RDC (Spitzer *et al*, 1978) and the Newcastle Diagnostic Index (NDI; Carney *et al*, 1965). This was done to increase differentiation, as the RDC have previously been criticised for their over-inclusiveness in the endogenous category. Length of current episode of illness and the presence of a concurrent personality disorder were not used as exclusion criteria.

Of 120 patients who were screened, 45 were rejected because they did not satisfy entry criteria or refused to be randomly allocated to treatment. After written consent, patients were randomly allocated to treatment according to a stratified model which included definite endogenous or non-endogenous depression (NDI), gender, age (three levels: 18–33; 34–49; 50–65), number of episodes (three levels: 2, 3 and 4 or more) and severity (HRSD <20 *v.* $\geqslant 20$). The stratification variables were chosen because they have been implicated in previous studies as predictors of response and/or of relapse/recurrence. Evaluators remained blind to treatment allocation.

Design

The study consisted of an acute treatment phase of 16 weeks and a follow-up phase of up to two years. Groups were defined according to treatment allocated during the acute treatment phase and during the follow-up phase. The 75 patients selected after screening were allocated according to the stratified design described above to three groups:

Group 1 (n=26): antidepressant medication (ADM) during acute treatment and antidepressant medication during follow-up (ADM→ADM).

Group 2 (n=22): antidepressant medication during acute phase of treatment and CT during follow-up (ADM→CT).

Group 3 (n=27): CT during acute phase of treatment and CT during follow-up (CT→CT).

Treatments

CT was administered by one of two extensively trained cognitive therapists (the authors). During the acute phase, CT typically took place once a week, with a minority of patients being seen twice a week at the outset of treatment. Patients maintained on CT saw their therapist three times in the first month, twice in the second month and monthly thereafter. For patients on ADM, consultants in charge or general practitioners (GPs) were free to prescribe any antidepressant of their choice or to switch medication during treatment, as long as the antidepressant was prescribed at or above therapeutic doses during the acute phase of treatment (the equivalent of 100 mg daily of amitriptyline for tricyclics, 45 mg daily of phenelzine for monoamine oxidase inhibitors and 20 mg daily of fluoxetine for selective serotonin reuptake inhibitors (SSRIs)). Patients receiving medication were seen at out-patient clinics by their consultants or registrars or by their GPs about every three weeks for half an hour or more to monitor their medication and for general support. During the follow-up phase, the maintenance dose of medication was left to the discretion of the doctor in charge, but again had to be above recognised maintenance doses (50 mg/day amitriptyline for tricyclics, 30 mg/day phenelzine for monoamine oxidase inhibitors and 20 mg/day fluoxetine for SSRIs).

Immediately after random allocation, patients completed a series of self-rating mood questionnaires and a number of measures of cognitive style, attitudes and personality. Treatment then started and ratings on all measures were repeated every four weeks during acute treatment and every four months during the follow-up phase. Thus, during the treatment phase, ratings were obtained at baseline, weeks 4, 8, 12 and 16. During the two-year follow-up, patients were rated at 4, 8, 12, 16, 20 and 24 months.

In this report, outcome data at and over 16 weeks and follow-up data at and over two years will be considered. Only measures relating to level of depression will be analysed, that is the HRSD and Beck Depression Inventory (BDI; Beck *et al*, 1961).

Statistical analysis

The data were analysed using analyses of covariance and repeated analyses of covariance, with baseline depression level and baseline variables which differentiated treatment groups as covariates. Where there were missing data with obtained data on either side, scores were prorated as the average of the two proximate scores.

Analyses for completers were complemented with end-point analyses to increase the power of analyses and to ensure that the results were representative of the originally randomly allocated groups. In the end-point analyses of acute treatment response, scores from the latest assessment completed after four weeks (inclusive) were carried forward and used as end-of-treatment data. In end-point analyses of follow-up results, patients' scores from the latest assessment completed after the end of acute treatment (inclusive) were considered as their final follow-up response. Categorical data were analysed using χ^2 tests.

Although Group 2 (ADM→CT) essentially had the same treatment as Group 1 (ADM→ADM) during the acute phase and the same treatment as Group 3 (CT→CT) during follow-up, the main analyses did not use collapsed groups, as it was considered that Group 2 may differ in important ways from the other groups, because of the switch in treatment to which they had given their consent. Subsidiary analyses did use collapsed groups to increase statistical power.

RESULTS

Patient's characteristics

Table 1 shows the baseline characteristics of all patients allocated to treatment. The three groups were well matched on all variables. The only difference was in healthiest level of functioning in the previous five years, as assessed on the SADS. Group 2 (ADM→CT) showed the lowest level of functioning, although a large proportion of all three groups had experienced significant levels of dysfunction in the five years prior to assessment (ratings $\geqslant 2$).

Table 1 Baseline characteristics of all patients allocated; mean (s.d.)

	Group 1 ADM→ADM	Group 2 ADM→CT	Group 3 CT→CT	P
n	26	22	27	
Gender (F/M)	17/9	17/5	14/13	NS
Age	40.1 (12.7)	37.8 (13.1)	39.6 (12.0)	NS
Socio-economic status (1+2/3+4+5)	12/14	13/9	16/11	NS
No. of episodes	3.2 (2.2)	4.1 (3.4)	3.0 (1.4)	NS
Range	2–11	2–15	2–8	
No. of hospitalisations	0.8 (2.3)	0.7 (0.9)	0.4 (0.7)	NS
Range	0–10	0–3	0–2	
Duration of index episode[1]	6.9 (1.3)	7.0 (1.4)	6.8 (1.4)	NS
<3 months/3 months–1 year/>1 year	6/14/6	5/11/6	8/13/6	
Healthiest level of functioning in previous 5 years[2]	1.5 (0.7)	2.0 (0.9)	1.4 (0.6)	0.02
No or minimal symptoms/symptoms present at best point	15/11	8/14	17/10	
Non-endogenous/endogenous (RDC)	10/16	12/10	16/11	NS
Non-endogenous/endogenous (NDI)	22/4	19/3	23/4	NS
Suicidal attempts	0.9 (1.9)	0.4 (0.7)	0.4 (0.7)	NS
Range	0–9	0–2	0–3	
BDI	29.9 (8.8)	28.9 (9.3)	27.7 (9.1)	NS
HRSD	20.2 (4.4)	20.8 (3.9)	19.9 (4.3)	NS

1. SADS ratings: 6=2 months–<3 months; 7=3 months–<1 year; 8=1 year–<2 years; 9=≥2 years.
2. SADS ratings: 1=absent or minimal symptoms; 2=slight impairment of functioning; 3=some mild symptoms; 4=moderate symptoms or general functioning with some difficulty; 5=serious symptoms or impairment; 6=major impairment in several areas.
RDC, Research Diagnostic Criteria; NDI, Newcastle Diagnostic Index; BDI, Beck Depression Inventory; HRSD, Hamilton Rating Scale for Depression.

Typically, the three groups showed moderate to severe levels of depression on both the HRSD and the BDI, and were experiencing their third episode of depression, although the range went up to 15 episodes. While 37 of the 75 patients showed a definite endogenous pattern of symptoms on the RDC, only 11 satisfied criteria for endogeneity on the NDI. These proportions are significantly different, confirming the stringency of the NDI (χ^2=20.7; d.f.=1; $P<0.0001$). The average number of previous hospitalisations was less than one, with a range of 0–10.

In the following analyses of variance, initial level of depression (although not differentiating the groups, but because of its known effects on response) and highest level of functioning in the previous 5 years, because this did differentiate the groups, were used as covariates.

Comparison of completers and non-completers

Of the 75 patients allocated to treatment, six dropped out before the first assessment at four weeks after treatment had started (one from Group 1, two from Group 2, and three from Group 3); two further patients from Group 1 were prescribed inadequate doses of medication. All these patients (n=8) were not included in further analyses. Of the 67 patients, five further patients (two from Group 1, three from Group 2) dropped out from treatment between four weeks and the end of the acute treatment phase. These patients were included in end-point analyses only, which thus included 23 in Group 1, 20 in Group 2 and 24 in Group 3. The final samples which completed 16 weeks of treatment (completers) were 21, 17 and 24, respectively, 62 patients altogether.

Comparison of the proportion of drop-outs and completers from the three treatment groups was not significant (χ^2=1.27; d.f.=2; NS). The 13 patients (17%) who dropped out at various points of the 16-weeks treatment period were compared with the 62 completers. There was no difference on any of the variables listed in Table 1, with only age showing a trend towards significance, non-completers (mean age 45.3 (s.d. 12.2)) being somewhat older than completers (mean age 38.2 (s.d. 12.3), P=0.08). The loss of 17% of the original sample is low compared with other studies (43% in Hollon *et al*, 1992; 35% in Elkin *et al*, 1989).

Outcome at 16 weeks

End-point data

Table 2 shows the unadjusted means of the three treatment groups and results of the repeated analyses of covariance (n=67) over the 16 weeks of acute treatment for end-point data, with baseline depression level and healthiest level of functioning in the five years preceding entry into the study as covariates.

It can be seen from Table 2 that all three treatment groups showed a highly significant improvement over time ($P<0.0001$) on both the HRSD and the BDI, but they did

Table 2 Outcome at 16 weeks for end-point data (unadjusted means (s.d.)) and *F*s for repeated analyses of covariance (baseline depression and healthiest level of functioning in previous five years as covariates)

	n	Baseline	Week 4	Week 8	Week 12	Week 16
HRSD						
Group 1	23	20.3 (4.3)	16.5 (5.8)	14.8 (6.9)	13.1 (7.1)	11.4 (7.3)
Group 2	20	20.6 (4.0)	15.8 (4.0)	15.7 (6.4)	14.3 (6.4)	13.3 (7.0)
Group 3	24	19.2 (3.6)	15.6 (6.6)	14.8 (6.8)	11.3 (6.9)	10.7 (7.6)
BDI						
Group 1	23	29.6 (9.1)	25.3 (13.1)	22.3 (14.3)	21.3 (15.0)	20.3 (15.4)
Group 2	20	28.8 (9.6)	24.7 (9.6)	24.4 (11.6)	22.7 (11.8)	22.8 (13.1)
Group 3	24	27.3 (8.7)	23.7 (9.6)	21.9 (12.9)	18.7 (13.9)	19.0 (12.5)

HRSD: F (treatment)=0.01, d.f.=2, 62, NS; F (time)=41.03, d.f.=4, 255, $P<0.0001$; F (treatment × time)=0.51, d.f.=8, 255, NS.
BDI: F (treatment)=0.04, d.f.=2, 61, NS; F (time)=15.36, d.f.=4, 255, $P<0.0001$; F (treatment × time)=0.42, d.f.=8, 251, NS.
Group 1: ADM→ADM; Group 2: ADM→CT; Group 3: CT→CT.
HRSD, Hamilton Rating Scale for Depression; BDI, Beck Depression Inventory.

not differ significantly and there was no interaction between treatment and time, indicating that the pattern of improvement was similar across groups. The order of response is CT→CT, ADM→ADM and ADM→CT on both severity measures. No significant difference emerged when the two medication groups were collapsed and compared with the CT group.

Completers' data

Analyses of pre- and post-treatment data for completers (n=62), again using baseline severity and highest level of functioning as covariates, also showed no significant difference between treatment groups. The pre- and post-treatment mean (s.d.) scores on the HRSD for the ADM→ADM, ADM→CT and CT→CT groups consecutively were: 20.6 (4.4) and 11.4 (7.6); 20.1 (3.5) and 11.8 (6.3); 20.0 (4.3) and 10.5 (7.2). The corresponding scores on the BDI for the three groups consecutively were: 30.1 (9.5) and 21.1 (16.1); 27.4 (7.1) and 19.9 (10.9); 27.7 (9.1) and 18.7 (12.2). None of the cross-sectional comparisons at evaluation points approached significance and no significant differences were obtained when the two medication groups were collapsed and compared with the CT group.

Categorical comparisons

Full response was defined as a score of ⩽6 on the HRSD, following criteria adapted in the two reference studies, and partial response as a score of 7–14. The HRSD was used rather than the BDI or a combination of the two measures because the HRSD is generally considered to be a more reliable measure of severity (Prusoff *et al*, 1972) and other studies, for example the reference studies of Elkin *et al* (1989) and Hollon *et al* (1992), have used the HRSD for determining end-of-treatment response. Considering only those patients who completed treatment, 24% reached full remission on antidepressants from both Group 1 and Group 2 and 33% on CT (Group 3). Partial response was achieved by 43, 41 and 33% and non-response was similar at 33, 35 and 35% in Groups 1, 2 and 3, respectively. The proportions of responders, partial responders and non-responders in the three treatment groups did not differ significantly (χ^2=0.94; d.f.=4). Collapsing the two groups on antidepressant medication during acute treatment produced the same lack of differentiation (χ^2=0.76; d.f.=2).

Comparison of more and less severely depressed

To investigate the relative effect of ADM and CT on the more (HRSD⩾20) or less severely depressed (HRSD<20), secondary analyses were done on completers. For the more severely depressed, the analysis of covariance on end of treatment scores (with baseline depression covaried) revealed an F of 0.16, d.f.=2, 24; NS. For the less severely depressed, F was 1.03, d.f.=2, 30; NS. Thus, neither treatment was shown to be superior for the less or more severely depressed patients.

Follow-up

End-point data

Table 3 shows the interpolated and end-point data at the end of treatment and for four-monthly assessments during the 24 months of follow-up.

On the HRSD, the repeated analysis of covariance (baseline level of depression and healthiest level of functioning in previous five years as covariates) showed no significant difference between treatments (F=0.31; d.f.=2, 55; NS), and no interaction between treatment and time (F=0.63; d.f.=12, 341; NS), but a highly significant difference for time (F=5.15; d.f.=2, 341; P<0.0001). Thus all patients continued to improve over the follow-up period. On the BDI, again the three treatment groups were not differentiated (F=0.72; d.f.=2, 53; NS), patients kept improving significantly over time (F=4.51; d.f.=6, 329; P<0.001) and there was a trend for an interaction between treatment and time (F=1.61; d.f.=12, 329; P<0.08). This trend appears to reflect a more steady improvement in the two CT groups (Group 2 and 3) from 12 months onwards.

In the subsidiary analyses, when the two groups were collapsed, no significant differences emerged.

Completers' data

It can be seen from Table 4 that from 12 months onwards, there was a dwindling number of patients, with 20 patients in all remaining at 24 months. This was due primarily to the fact that a large number of patients were recruited too late to complete the two-year follow-up within the study period. Cross-sectional analyses of covariance at each assessment point (baseline level of depression and healthiest level of functioning in previous years as covariates) showed that on both the HRSD and the BDI, the group treated with and followed up on CT did marginally better over time than the other two groups. However, there were no significant differences between treatments at any point in time on the HRSD and only one trend for a difference on the BDI at 20 months, with the two CT groups doing better than the medication group.

Again, in subsidiary analyses, when the two CT groups were collapsed, there were no significant differences.

Categorical analyses

Chi-squared tests were carried out for the data at 12 months only, as enough patients were available in the various cells for this type of analysis. The HRSD was used to define recovery, combining the fully recovered and the partially recovered (HRSD⩽14), as

Table 3 Follow-up over 24 months for end-point data (unadjusted means (s.d.)) and Fs for repeated analyses of covariance (baseline depression and healthiest level of functioning over previous five years as covariates)

	n	End of treatment	4 months	8 months	12 months	16 months	20 months	24 months
HRSD								
Group 1	20	10.6 (6.8)	10.7 (7.8)	9.4 (7.4)	10.4 (7.2)	9.4 (7.1)	9.6 (7.1)	9.3 (7.2)
Group 2	17	11.8 (6.3)	10.5 (4.2)	10.3 (7.3)	9.2 (7.7)	9.1 (6.6)	7.9 (6.3)	8.6 (5.6)
Group 3	23	10.1 (7.1)	10.0 (6.7)	8.7 (7.6)	7.8 (7.3)	7.1 (7.0)	6.3 (6.5)	6.6 (7.0)
BDI								
Group 1	20	19.7 (14.2)	18.4 (14.1)	16.8 (14.3)	19.6 (15.2)	19.3 (16.2)	19.9 (14.2)	18.1 (13.1)
Group 2	16	20.4 (11.1)	20.6 (7.2)	14.3 (9.5)	14.9 (12.4)	12.9 (9.5)	11.5 (8.5)	14.2 (9.9)
Group 3	22	18.2 (12.2)	19.8 (12.3)	17.1 (14.7)	14.8 (10.5)	14.3 (12.0)	12.7 (11.0)	13.3 (10.7)

HRSD: F (treatment)=0.31, d.f.=2, 55 NS; F (time)=5.15, d.f.=6, 341, P<0.0001; F (treatment×time)=0.63, d.f.=12, 341, NS. BDI: F (treatment)=0.72, d.f.=2, 53, NS; F (time) 4, 51, d.f.=6, 329, P<0.001; F (treatment×time) 1.61, d.f.=12, 329, NS, P<0.08. Groups as in Table 2.

Table 4 Outcome over follow-up for completer sample at each assessment point (unadjusted means (s.d.)) and *F*s for adjusted means at each assessment point (baseline depression and healthiest level of functioning as covariates)

	4 months	8 months	12 months	16 months	20 months	24 months
HRSD						
Group 1	10.2 (7.8)	8.1 (7.2)	9.2 (7.1)	7.4 (6.5)	7.4 (5.9)	7.2 (8.3)
n	16	17	17	12	7	5
Group 2	9.9 (3.9)	9.7 (7.5)	9.8 (8.3)	9.6 (6.5)	8.2 (6.5)	11.5 (4.9)
n	15	15	12	11	10	6
Group 3	9.5 (6.2)	7.6 (7.0)	6.5 (6.7)	6.4 (6.9)	5.6 (5.3)	6.8 (7.3)
n	19	19	18	13	12	9
F (d.f.)	0.68 (d.f.=2, 45), NS	0.23 (d.f.=2, 46), NS	0.44 (d.f.=2, 42), NS	0.12 (d.f.=2, 31), NS	0.57 (d.f.=2, 24), NS	0.01 (d.f.=2, 15), NS
BDI						
Group 1	19.5 (15.0)	16.7 (15.3)	20.0 (16.3)	17.4 (16.6)	20.2 (13.9)	13.2 (11.4)
n	16	17	17	11	8	5
Group 2	19.8 (7.6)	14.1 (9.4)	17.2 (13.6)	13.1 (10.0)	11.3 (8.0)	19.0 (10.5)
n	15	15	13	11	10	6
Group 3	18.5 (11.1)	15.4 (14.9)	12.5 (9.2)	12.3 (11.8)	10.7 (9.0)	13.1 (8.9)
n	17	18	17	13	12	8
F (d.f.)	0.08 (d.f.=2, 43), NS	0.32 (d.f.=2, 45), NS	0.99 (d.f.=2, 42), NS	0.69 (d.f.=2, 30), NS	2.67 (d.f.=2, 25), $P<0.09$	0.34 (d.f.=2, 14), NS

Groups as in Table 2.

compared with the non-recovered (HRSD > 14). This principle was adopted, as a cut-off of 14 on the HRSD is often used to define depression at entry into outcome studies (Elkin *et al*, 1989; Hollon *et al*, 1992) and previous studies have used a cut-off of 14 or 15 to define relapse (for example Thase *et al*, 1992).

Excluding patients who were not recovered at the end of acute treatment, the proportions of relapses were 31% for patients treated and maintained on medication (four out of 13), 36% for patients who switched from medication to CT during follow-up (five out of 14) and 24% for the CT-treated and maintained group (four out of 17). When patients who were receiving CT during follow-up were considered together to increase numbers for analysis, the proportion of the CT-maintained group who relapsed was 26% and the proportion of the medication maintained groups was 31% ($\chi^2=0.01$; d.f.=1).

DISCUSSION

The main findings of the study did not disconfirm the null hypotheses defined in the introduction. In out-patients selected for recurrent major depression, unipolar subtype, cognitive therapy was as effective as medication for the treatment of the acute episode. The non-significant ordering of efficacy in our analyses was cognitive therapy (group followed on cognitive therapy), antidepressants (group followed on antidepressants) and antidepressants (group followed on cognitive therapy).

In long-term outcome, maintenance cognitive therapy was as effective as maintenance medication, whether cognitive therapy followed acute treatment with cognitive therapy or with medication. Again, the non-significant ordering of long-term efficacy in all analyses was cognitive therapy (following acute treatment with cognitive therapy), antidepressants (following acute treatment with antidepressants), but the cognitive therapy (following acute treatment with antidepressants) group did not consistently do worse than the group treated and maintained on medication. The consistent lower response during the acute phase of treatment of the group treated with medication and followed-up on cognitive therapy, although non-significant, is intriguing (at least in theory, groups 1 and 2 were receiving the same treatment and their responses should have been similar). It may reflect a mental set that somehow interfered with response to medication. Many patients in this group told us that they were eager to stop medication and start cognitive therapy during their follow-up treatment.

Each of these findings will be discussed in turn, within the context of the sample characteristics, the two reference studies and other relevant outcome studies. It must be borne in mind that in view of the small numbers in the samples, particularly in the maintenance data, the lack of differentiation between groups may be due to a type II error.

The sample

The patients selected in this study were of a comparable level of depression with those in Elkin *et al* (1989) and Hollon *et al* (1992), although the cut-off point on the HRSD at entry was higher (score of 16 compared with score of 14). They differed in that they were selected for the recurrent subtype of depression. The mean number of episodes ranged from three to four, with some patients having suffered over 10 episodes, whereas over a third of the sample in Elkin *et al* and over a quarter in Hollon *et al* had suffered no previous episode of depression. Moreover, although the duration of the current episode was typically three months to one year (51%), 24% of the sample had a current episode of over one year and only 25% had a current episode of less than three months. Thus, this sample was considerably more chronic than that of Hollon *et al* (1992) for whom 79% had a current episode of less than six months' duration, but fairly similar to Elkin's *et al* (1989) for whom 35% had a current episode of over one year duration.

The exclusion criteria used in the different studies must be considered. To increase the number of potential recruits, with time and resources constraints, unlike the two reference studies, we did not exclude any type of personality disorders, concurrent secondary psychiatric disorders, such as panic disorder, generalised anxiety, obsessive–compulsive disorder, phobic disorder and secondary alcoholism. Both the presence of personality disorders and secondary pathology have been found to be negative predictors of response.

The percentage of patients who dropped-out of treatment, including those excluded by study staff, was extremely low at 17% compared with 32% and 35% in Elkin *et al* (1989) and Hollon *et al* (1992), respectively.

Short-term outcome

The equivalence of effect on the HRSD and the BDI between cognitive therapy and antidepressant medication was reported by Hollon *et al* (1992) and Elkin *et al* (1989)

for both completers and end-point analyses. Previous studies comparing cognitive therapy and antidepressant medication singly have shown mixed results, with CT showing more efficacy in some studies and the two treatment methods being of equal efficacy in others. Comparison of the proportions of recovered patients in this study with those in the two reference papers, using the same strict criterion of $\leqslant 6$ on the HRSD for completer samples, indicates that this study achieved lower levels of asymptomatic patients. We found that 24% of patients treated with medication (pooled samples) and 33% of patients treated with CT were fully recovered, compared with 57 and 51% in Elkin *et al* (1989) and 53 and 50% in Hollon *et al* (1992). Neither study reports non-response rates (HRSD $\geqslant 14$), which, in this study, were 34% for medication (pooled groups) and 33% for cognitive therapy. These figures are comparable with previous findings in non-response (for example, 32% for the combination of imipramine and psychotherapy; Frank *et al*, 1990). None the less, some explanations are necessary for the lower full response rate in this study. This might be attributed to the difference in patient samples, the difference in treatment delivery, or the difference in the rating style of objective raters across studies.

The difference in the samples has already been commented on. Ours were National Health Service patients with recurrent depression, relatively chronic and well distributed across socio-economic classes. The NIMH study reported a primarily higher educated sample, whereas Hollon *et al* (1992) described a primarily lower middle-class sample. There is no indication to date that socio-economic class relates to response. There is, on the other hand, some indication that recurrent depression responds less well than first-time depression (Thase, 1992) and that chronicity is a negative predictor of response (Keller *et al*, 1986).

The difference in treatment delivery cannot be excluded. In the case of medication, any recognised antidepressant at or above recognised therapeutic doses was accepted in our study, while both the American studies, as most other comparative studies, opted for one antidepressant, namely the tricyclic imipramine hydrochloride up to 300 mg daily (Hollon *et al*, 1992) and 185 mg daily (Elkin *et al*, 1989). The rationale for adopting our strategy was to mimic good clinical practice as much as possible. We left the choice of the antidepressant and the option to change medication during treatment to the medical prescriber because individual patients might not have responded to one or other antidepressant in a previous episode of depression. In the case of cognitive therapy, the two therapists were trained cognitive therapists who followed the same model of therapy as the therapists in the two reference studies did (Beck *et al*, 1979) and whose therapeutic allegiance was cognitive therapy. Mutual supervision during the course of the study took place on a weekly basis, although no formal monitoring of therapy sessions was made. There is no *prima facie* reason to believe that the quality of the therapy was not adequate, but the settings in which therapy was delivered may well have differed. The setting in which therapy is delivered has been shown to have an effect on training and on patients' behaviours (McDonald, 1991). Length of treatment (16 weeks, average 15 sessions for completers) did not differ markedly from that in the NIMH study (16 weeks, average 16.2 sessions for completers), or that in the Hollon *et al* study (12 weeks, average 14.9 sessions for completers).

There was no evidence in our study that level of severity defined as HRS scores $\geqslant 20$ differentiated responders to cognitive therapy or to antidepressants, disconfirming the NIMH findings (Elkin *et al*, 1989, 1995) and supporting Hollon *et al* (1992). Thase *et al* (1991) and McLean & Taylor (1992).

Long-term outcome

End-point analyses indicated that, over two years' controlled follow-up, patients in all groups continued to improve highly significantly over time on both outcome measures, while there was no significant difference between treatments or interaction between time and treatment. Thus all three treatment groups showed the same pattern of continuous improvement over two years. That short-term treatment may not be sufficient in some cases has been noted both for cognitive therapy (Thase *et al*, 1992) and for antidepressants (Frank *et al*, 1990).

For patients with complete data, cross-sectional analyses indicated no difference between treatments at any point during follow-up, with only a trend on the BDI at 20 months in favour of cognitive therapy. Thus maintenance cognitive therapy was as effective as maintenance medication.

An investigation of what happened to individuals over the follow-up period, rather than relying on group means which mask the variability of response, was carried out only at 12 months because of numbers left in the study. The relapse/recurrence rates for patients who had recovered at the end of treatment showed a slight, but non-significant, advantage for patients treated and maintained on cognitive therapy. Shea *et al* (1992) found a relapse rate at 12 months of 28% (compared with 31% in our study) in their medication plus clinical management group, and 9% in their cognitive therapy group (compared with 24% in the present study). Evans *et al* (1992), at 24 months, found a relapse rate of 32% in the group treated with medication and maintained on medication for six months, 50% for patients treated with medication and not maintained on medication, and 21% for patients treated with cognitive therapy.

Thus, our drug-treated and drug-maintained group did slightly worse than the drug-treated group of Shea *et al*, better than the Evans *et al* (non-drug maintained) group at two years and about the same as their partially maintained group. In relation to studies of maintenance medication in recurrent depression specifically, the recurrence rate of 31% in the drug-maintained group compares favourably with 45% found by Glen *et al* (1984) at one year in patients with recurrent depression maintained on lithium or amitriptyline, and 54% found by Georgotas *et al* (1989) at one year in patients maintained on nortriptyline. However, with high maintenance dosages of over 200 mg imipramine daily, Frank *et al* (1990) found a recurrence rate of 17.9%.

The relapse rate in our cognitive therapy-treated and maintained group is considerably worse than that of Shea *et al* at one year. It is slightly higher than that reported by Evans *et al* at two years, by Blackburn *et al* (1986) at two years (23%) and by Simons *et al* (1986) at one year (20%). This difference is likely to be attributable to sample characteristics, ours being selected for recurrent depression. Again, the group that switched over from antidepressant to cognitive therapy did worse than the other two groups. However, the proportion of relapses in the different groups did not differ significantly and as the mean level of depression at different assessment points in the three groups did not differ significantly, the tentative conclusion is that for patients who do not wish to or cannot be maintained on antidepressant medication, infrequent cognitive therapy is a viable maintenance treatment after acute treatment with medication.

ACKNOWLEDGEMENTS

This research was carried out in the Department of Psychiatry, University of Edinburgh, with a grant from the Scottish Home and Health Department. Grateful thanks to Angela Wilson and Michèle Hipwell for their help in data collection; to Dr Delia Wakelin for running the computer programs and to Katharina Reichelt for the word processing.

REFERENCES

Beck, A. T., Ward, C. H., Mendelson, M., *et al* (1961) An inventory for measuring depression. *Archives of General Psychiatry*, **4**, 561–571.

—, Rush, A. J., Shaw, B. F., *et al* (1979) *Cognitive Therapy of Depression*. New York: Guilford.

Blackburn, I.-M., Eunson, K. M. & Bishop, S. (1986) A two-year naturalistic follow-up of depressed patients treated with cognitive therapy, pharmacotherapy and a combination of both. *Journal of Affective Disorders*, **10**, 67–75.

— & Twaddle, V. (1996) *Cognitive Therapy in Action. A Practitioner's Casebook*. London: Souvenir Press.

Carney, M. W. P., Roth, M. & Garside, R. F. (1965) The diagnosis of depressive syndromes and the prediction of ECT response. *British Journal of Psychiatry*, **111**, 659–674.

Elkin, I., Shea, M. T., Watkins, J. T., *et al* (1989) National Institute of Mental Health Treatment of Depression Collaborative Research Program. General effectiveness of treatments. *Archives of General Psychiatry*, **46**, 971–982.

—, —, —, *et al* (1995) Initial severity and differential treatment outcome in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology*, **63**, 841–847.

Endicott, J., Spitzer, R. L., Fleiss, J. Z., *et al* (1976) The Global Assessment Scale: a procedure measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry*, **33**, 766–771.

— & — (1978) A diagnostic interview: The schedule for affective disorders and schizophrenia. *Archives of General Psychiatry*, **35**, 837–844

Evans, M. D., Hollon, S. D., DeRubeis, R. J., *et al* (1992) Differential relapse following cognitive therapy and pharmacotherapy for depression. *Archives of General Psychiatry*, **49**, 802–808.

Frank, E., Kupfer, D. J., Perel, J. M., *et al* (1990) Three-year outcome for maintenance therapies in recurrent depression. *Archives of General Psychiatry*, **47**, 1093–1099.

Georgotas, A., McCue, R. E. & Copper, T. B. (1989) A placebo-controlled comparison of nostriptyline and phenelzine in maintenance therapy of elderly depressed patients. *Archives of General Psychiatry*, **46**, 783–786.

Glen, A. I. M., Johnson, A. L. & Shepherd, M. (1984) Continuation therapy with lithium and amitriptyline in unipolar depressive illness: a randomized, double blind, controlled trial. *Psychological Medicine*, **14**, 37–50.

Greenhouse, J. B., Stangl, D., Kupfer, D. J., *et al* (1991) Methodologic issues in maintenance therapy clinical trials. *Archives of General Psychiatry*, **48**, 313–318.

CLINICAL IMPLICATIONS

- For a group of depressed patients with recurrent major depression, short-term cognitive therapy was as efficacious as antidepressants prescribed at therapeutic doses.
- Maintenance cognitive therapy over two years was as effective as maintenance medication.
- Maintenance cognitive therapy after acute treatment with medication can be a viable alternative to maintenance medication.

LIMITATIONS

- The results are based on relatively small numbers.
- No objective measure of compliance with medication was used.
- The number of recurring episodes varied widely. More-restricted criteria of recurrence may be more informative.

I.-M. BLACKBURN, Newcastle Cognitive and Behavioural Therapies Centre; R. G. MOORE, Department of Psychiatry, University of Cambridge, Addenbrooke's Hospital (Box 189), Hills Road, Cambridge CB2 2QQ

Correspondence: Professor I.-M. Blackburn, Newcastle Cognitive and Behavioural Therapies Centre, Collingwood Clinic, St Nicholas Hospital, Gosforth, Newcastle upon Tyne NE3 3XT

(First received 5 November 1997, final revision 10 April 1997, accepted 14 April 1997)

Hamilton, M. A. (1960) A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry*, **23**, 56–62.

Hollon, S. D., DeRubeis, R. J., Evans, M. D., *et al* (1992) Cognitive therapy and pharmacotherapy for depression singly and in combination. *Archives of General Psychiatry*, **49**, 774–781.

Keller, M. B., Lavori, P. B., Klerman, G. L., *et al* (1986) The persistent risk of chronicity in recurrent episodes of non bipolar major depressive disorder. A prospective follow-up. *American Journal of Psychiatry*, **143**, 24–28.

McDonald, R. M. (1991) Assessment of organizational context: A missing component in evaluations of training programs. *Evaluation and Program Planning*, **14**, 273–277.

McLean, P. D. & Taylor, S. (1992) Severity of unipolar depression and choice of treatment. *Behaviour Research and Therapy*, **30**, 443–451.

Prusoff, B. A., Klerman, G. L. & Paykel, E. S. (1972) Concordance between clinical assessments and patients' self-report in depression. *Archives of General Psychiatry*, **26**, 546–552.

Rush, A. J., Beck, A. T., Kovacs, M., *et al* (1977) Comparative efficiency of cognitive therapy versus pharmacotherapy in outpatient depression. *Cognitive Therapy and Research*, **1**, 17–37.

Shea, M. T., Elkin, I., Imber, S. D., *et al* (1992) Course of depressive symptoms over follow-up. Findings from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Archives of General Psychiatry*, **49**, 782–787.

Simons, A. D., Murphy, G. E., Levine, J. L., *et al* (1986) Cognitive therapy and pharmacotherapy for depression: sustained improvement over one year. *Archives of General Psychiatry*, **43**, 43–48.

Spitzer, R. L., Endicott, J. & Robins, E. (1978) Research diagnostic criteria: Rationale and reliability. *Archives of General Psychiatry*, **35**, 773–782.

Thase, M. E. (1992) Long-term treatments of recurrent depressive disorders. *Journal of Clinical Psychiatry*, **53** (suppl.), 32–44.

—, Simons, A. D., Cahalane, J., *et al* (1991) Severity of depression and response to cognitive behavior therapy. *American Journal of Psychiatry*, **148**, 784–789.

—, —, McGeary, J., *et al* (1992) Relapse after cognitive behavior therapy of depression: potential implications for longer courses of treatment. *American Journal of Psychiatry*, **149**, 1046–1052.

Controlled acute and follow-up trial of cognitive therapy and pharmacotherapy in out-patients with recurrent depression.

I M Blackburn and R G Moore

BJP 1997, 171:328-334.
Access the most recent version at DOI: 10.1192/bjp.171.4.328